A FRAMEWORK OF PASSIVE-ACTIVE-CONSTRUCTIVE STUDY TECHNIQUES:

A DIVERGENCE BETWEEN ASSIGNED AND REPORTED BEHAVIORS

A Thesis
Presented to
The Academic Faculty

by

Keith R Bujak

In Partial Fulfillment
of the Requirements for the Degree
Masters of Psychology in the
School of Psychology

Georgia Institute of Technology
December, 2010

A FRAMEWORK OF PASSIVE-ACTIVE-CONSTRUCTIVE STUDY TECHNIQUES:

A DIVERGENCE BETWEEN ASSIGNED AND REPORTED BEHAVIORS

Approved by:

Dr. Richard Catrambone, Advisor
School of Psychology
Georgia Institute of Technology

Dr. Gregory Corso
School of Psychology
Georgia Institute of Technology

Dr. Mark Guzdial
School of Interactive Computing
Georgia Institute of Technology

Date Approved: 07/26/10

ACKNOWLEDGEMENTS

I wish to thank the members of my committee for their thoughtful insights and challenging questions. I wish to thank my lab mates for their helpful feedback, especially as I was learning to focus my research questions.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

SUMMARY

An educational framework proposed by Chi (2009) aims to link overt study activities with outcomes via the underlying cognitive processes experienced by learners. Activities are classified along a continuum of passive, active, and constructive[1]. Overt activities—such as reading, highlighting, and self-explaining—are grouped according to the hypothesized cognitive processes they engage. The framework posits that constructive activities yield the best learning, passive yield the poorest, and active is somewhere in between. Although these hypotheses are not supported by this experiment, there is evidence to suggest that college students employ study techniques that go beyond what they are asked to do. Also, the content of the text to be studied is potentially an important factor for determining the type of studying learners do regardless of what they are asked to do. In sum, although the framework is supported by many other studies, there might be additional variables that need to be considered when implementing this framework.

---

[1] Chi's original framework includes a fourth level, interactive, in which pairs of students learn from each other. The focus of this study is individual learning, therefore the fourth level is not included here.

CHAPTER 1:

INTRODUCTION

1.1 Theoretical Framework

Chi (2009) proposed a framework to link overt study activities with the underlying cognitive processes that support learner outcomes. The framework considers four levels of overt learning activities in which students can engage and speculates on the potential underlying cognitive processes exercised by these activities. This research explores just the first three activity levels of the framework: passive, active, and constructive. The fourth activity level, interactive, describes how two learners can share knowledge for the benefit of one or both. This activity level will be left for future studies as the present study is focused on individual learning. For each of the three activity levels currently under investigation, both the overt activities and underlying cognitive processes are described in more detail. Overt cognitive processes are defined as observable actions undertaken by the learner (e.g., highlighting, note-taking, self-explaining, generating questions). Underlying cognitive processes are defined as those theoretical psychological constructs (e.g., attending, integrating knowledge, restructuring mental models) engaged by a particular overt activity. Chi posists that each level subsumes the ones below it. Active learning requires the learner to experience both passive and active learning situations, gaining the benefits of both of those levels. Constructive situations would involve passive and active aspects, as well as activities unique to the constructive level.

Students might be asked, for example, to learn the text shown in Figure 1, which is an excerpt from a chapter of a popular undergraduate biology textbook (Campbell &

Reece, 2007). According to the framework, learners could engage with this material at one of three levels: passive, active, or constructive. This excerpt will be referenced throughout this manuscript when describing overt activities, assessment items, and other ideas.

> ***Ecological Niches***
>
> The sum of a species' use of the biotic and abiotic resources in its environment is called the species' ecological niche. American ecologist Eugene Odum used the following analogy to explain the niche concept: If an organism's habitat is its "address;' the niche is the organism's "profession:' Put another way, an organism's niche is its ecological role—how it "fits into" an ecosystem. For example, the niche of a tropical tree lizard consists of, among many components, the temperature range it tolerates, the size of branches on which it perches, the time of day when it is active, and the sizes and kinds of insects it eats.
>
> We can use the niche concept to restate the principle of competitive exclusion: Two species cannot coexist permanently in a community if their niches are identical. However, ecologically similar species can coexist in a community if there are one or more significant differences in their niches. When competition between species with identical niches does not lead to local extinction of either species, it is generally because one species' niche becomes modified. In other words, evolution by natural selection can result in one of the species using a different set of resources. The differentiation of niches that enables similar species to coexist in a community is called resource partitioning (Figure 1). You can think of resource partitioning in a community as "the ghost of competition past"—the indirect evidence of earlier interspecific competition resolved by the evolution of niche differentiation.

Figure 1. Example of text to be read and learned.

1.1.1 Passive Learning: Overt Activities

Considered to be the most rudimentary level of learning, passive describes a learner who is superficially engaged in the learning experience. Overt activities are limited, particularly when compared to the other levels of the framework. The learner is considered to be a passive receptacle for information. Common overt activities at this level include listening, viewing, and reading. In the case of listening and viewing, the learner is provided with no opportunity to control the experience. Example activities include listening to a lecture or viewing a video. The learner is merely “along for the ride.” In regards to reading, the learner is primarily concerned with the task of reading itself (i.e., decoding and encoding words of the text). Although it is possible to read while engaging deep cognitive processes, reading in a passive manner assumes that the learner is not engaging in these deep activities, as they are hallmarks of the other levels in the framework.

1.1.2 Passive Learning: Cognitive Processes

Reading comprehension is expected to suffer when the learner reads passively. One concept that potentially explains this limited comprehension is “mind wandering.” While merely reading a mystery without focusing their attention on the story, readers failed to glean vital pieces of information from the text to create the appropriate connections in order to identify the villain (Smallwood, McSpadden, & Schooler, 2008). Based on Chi’s framework, however, it is likely that the learner is able to obtain some new information, although this information is unlikely to be integrated effectively with existing information in the learner’s memory. The learner will likely possess a scattering of new unconnected facts.

1.1.3 Active Learning: Overt Activities

Active learning is suggested to be the next most effective level of learning. This level is characterized by overt actions undertaken by the learner. The learner performs actions in the learning environment, such as searching, pointing, underlining, and selecting. The learner is not, however, creating anything that goes beyond what is explicitly presented. The learner is either moving various parts of his or her body or manipulating parts of the learning environment.

Two specific active activities are highlighting and taking copy-paste notes. In each of these cases, the learner is instructed to focus on the main ideas in each paragraph. The learner reads each of the paragraphs and decides which sentence is the most important (Igo, Bruning, & McCrudden, 2005). In a computer-based environment, learners in both cases select the appropriate text and then either click the highlight tool or copy the text and paste it in another location. Figure 2 shows the excerpt with the most important sentences, as determined by a hypothetical learner, highlighted. A learner tasked with taking copy-paste notes would take these same sentences and, instead of highlighting them, copy them and paste them into another document.

Ecological Niches

The sum of a species' use of the biotic and abiotic resources in its environment is called the species' ecological niche. American ecologist Eugene Odum used the following analogy to explain the niche concept: If an organism's habitat is its "address;' the niche is the organism's "profession:' Put another way, an organism's niche is its ecological role—how it "fits into" an ecosystem. For example, the niche of a tropical tree lizard consists of, among many components, the temperature range it tolerates, the size of branches on which it perches, the time of day when it is active, and the sizes and kinds of insects it eats.

We can use the niche concept to restate the principle of competitive exclusion: Two species cannot coexist permanently in a community if their niches are identical. However, ecologically similar species can coexist in a community if there are one or more significant differences in their niches. When competition between species with identical niches does not lead to local extinction of either species, it is generally because one species' niche becomes modified. In other words, evolution by natural selection can result in one of the species using a different set of resources. The differentiation of niches that enables similar species to coexist in a community is called resource partitioning (Figure 1). You can think of resource partitioning in a community as "the ghost of competition past"—the indirect evidence of earlier interspecific competition resolved by the evolution of niche differentiation.

Figure 2. Example text with main ideas highlighted based on the text from Figure 1.

1.1.4 Active Learning: Cognitive Processes

The cognitive processes hypothesized to underlie this level can be summarized as attending, which help to direct the learner's attention to critical parts of the material. Through this process of attending, the learner is able to activate existing knowledge; assimilate, activate, or store new information; and search existing knowledge. Taking the passively learned facts from the previously discussed level, it is likely that now the learner is able to better incorporate these new facts. Structure is given to the knowledge and gaps in understanding can be filled.

1.1.5 Constructive Learning: Overt Activities

Moving up a level in this framework of learning, constructive activities require the learner to produce outputs that go beyond what is explicitly presented. The learner engages in activities that support synthesizing new information. This goal is supported by encouraging the learner to make new connections in the material through activities such as concept mapping and predicting outcomes.

Two specific constructive activities are self-explaining and self-questioning. Self-explaining tasks the learner with making connections in the material through considering what is important and describing what the materials means to the learner. Self-questioning tasks the learner with developing questions about the materials. Questions are meant to be thoughtful and reflective, requiring answers that go beyond what is explicitly presented in the text. The learner must answer the questions, thus utilizing the knowledge gained from the material. Using the text from Figure 1, examples of self-explanations and self-questions are shown in Table 1.

Table 1
Example self-explanations and self-questions based on the text from Figure 1.

Constructive activity	Example response
Self-explanation	"Species requiring identical resources cannot live in exactly the same place. It is possible for the species to coexist temporarily, but something must change or else one species will become extinct. This form of competition will lead to adoptions of one or both species. It seems likely that adapting in this manner will lead to increased—but distributed—consumption of resources in the area."
Self-question	"What are the strengths and weaknesses, from an evolutionary standpoint, of ecological niches? Ecological niches enable greater species diversity by allowing species to coexist in similar, but not identical, areas. On the other hand, if there are more species in a given area, there is the chance that resources will be over-exploited."

1.1.6 Constructive Learning: Cognitive Processes

A variety of creation processes underlie this level of activity. Organizing existing and new information can help learners to create knowledge in a more meaningful way or to repair their own faulty knowledge. Because each level subsumes the ones below it, in order to reach these advanced levels of knowledge creation, the learner must also acquire at least some basic information in a passive manner (e.g., by reading), as well as attend to important information in an active manner.

1.1.7 Summary of Activities and Processes

Each of these three levels of activity presumably engages certain cognitive processes. Table 2 provides a description of the overt activities and proposed underlying cognitive processes for each of the three levels.

Table 2
Summary of activities and cognitive processes (adapted from Chi, 2009).

	Passive	Active	Constructive
Characteristics	None	Doing something physically	Producing outputs that contain ideas that go beyond the presented information
Overt activities	*Receiving Activities* • Read • Listen • View	*Engaging Activities* • Look, gaze, or fixate • Underline or highlight • Gesture or point • Paraphrase • Manipulate objects or tapes • Select • Repeat	*Self-construction Activities* • Explain or elaborate • Justify or provide reasons • Connect or link • Construct a concept map • Reflect or self-monitor • Plan and predict outcomes • Generate hypotheses
Cognitive processes	*Linguistic Processes* • Decode • Encode • Mind wander	*Attending Processes* • Activate existing knowledge • Assimilate, encode, or store new information • Search existing knowledge	*Creating Processes* • Infer new knowledge • Integrate new information with existing knowledge • Organize own knowledge for coherence • Repair own faulty knowledge • Restructure own knowledge

1.2 Statement of Problem

The current study represents the first known experiment to explicitly test the hypotheses of this framework simultaneously involving three levels of the framework (i.e., passive, active, constructive). Studies used to support the framework were conducted in a pairwise fashion (e.g., active versus passive, constructive versus constructive) or in a *post hoc* manner (e.g., participants were left to their own devices and outputs from the learning activity were categorized as passive, active, or constructive).

Second, multiple measures are used to determine compliance with each of the three levels. It is suggested in the original framework that students learning in one particular level (e.g., passive) might be *covertly* engaging in activities of another level (e.g., self explaining, which is a constructive activity). If a learner is tasked with the activities of one particular level but is, in fact, performing activities characteristic of another level, reinterpretation of their results might be necessary.

Third, detailed assessments are used to more clearly identify learning outcomes. By matching learning activities with learning outcomes, the underlying cognitive processes are likely to be more apparent. Typically, overall learning outcomes have been used to compare the different levels of learning activities. With overall assessment measures, it is not particularly clear what specific benefits learners are receiving from each of the activity levels. In this study, assessment items ranging from factual recall to near transfer are used. These measures hold the potential to elucidate the cognitive processes occurring during each of the three types of learning activities.

Fourth, both immediate and delayed tests are considered integral to learning. A general goal of education is to provide students with knowledge that persists after learning has concluded. Even if a particular activity results in superior performance immediately thereafter, if students lose the ability to demonstrate this mastery after even a short delay, the activity will have failed at achieving one of the basic goals of education.

Fifth, time-on-task is used explicitly as a dependent measure to explore the possibility that the learning differences due to activity level are not just a function of the activities performed. The framework posits that each level subsumes the ones below it. Without exposing some learners to more or different materials, learners in higher levels will inevitably spend more time with the materials in order to complete the activities. Perhaps it is simply this additional time that accounts for learning differences.

1.3 Literature Review

The studies reviewed below are the same as those used to originally support the framework (Chi, 2009). These studies are deemed by Chi to be ideal examples. For the purposes of designing this experiemnt, the studies were reviewed with attention focused on the five issues discussed above. The following section is organized in the same manner: simultaneous testing of levels, compliance with learning activities, detailed assessments, delayed testing, and learning time-on-task.

1.3.1 Simultaneous Testing

The original framework offers pairwise comparisons in support of its hypotheses, so this study represents the first known explicit attempt to compare more than two levels of activity simultaneously with the same learning materials used in each condition. Chi reviews these studies and assigns the conditions to particular levels of the framework. She then compares each of the studies in a pairwise fashion, both in terms of different activity levels (e.g., passive versus constructive) and in terms of equivalent levels (e.g., one type of active versus another type of active). Although each of the studies selected by Chi supports the framework, none explicitly manipulates more than two levels. The present study uses identical materials for all conditions, manipulated solely by the directions given to participants in order to elicit passive, active, and constructive activities.

1.3.2 Compliance with Activities

It is possible for a particular task to be altered in slight ways that might seem insignificant to the unassuming learner. These slight changes can cause the task to be reclassified along the continuum of passive-active-constructive. Consider, for example, note taking. Reading the notes of a lecture taken by a peer (e.g., Kiewra et al., 1991) can be considered passive, while taking verbatim notes (e.g., Trafton & Trickett, 2001) can be considered active, while taking conceptual notes and reviewing them (e.g., King, 1992) can be considered constructive. In fact, it has been found that when participants are provided with various note taking functions in a problem solving exercise without explicit directions on how to take notes, *post hoc* analyses of the notes found passive, active, and constructive components (Trafton & Trickett, 2001).

In some studies offered to support the framework, students spent extended periods of time outside the watchful gaze of the experimenters, during which they might have failed to engage in the required activities of their assigned condition. Or, they might have spontaneously engaged in activities that would classify them in a different level than the one to which they were assigned. During a mapping exercise in which students were to place stickers on a 2D map that corresponded to real-world objects, they spent time wandering about a park beyond the direct supervision of the experimenters (Kastens & Liben, 2007). Some participants were instructed to self-explain the placement of the stickers and others were not required to do so. It is likely that some of the more advanced students engaged in self-explaining strategies while attempting to properly place the stickers, making their activities constructive rather than passive. In another example, students from a class worked with an intelligent tutor to learn about the concepts of data normalization (Mitrovic, 2005). Participation in the study was optional and was assigned for homework. Out of twenty-nine students who initially participated, only six completed the experiment. Further, engagement with the tutoring system ranged from three to 653 minutes. Given this high attrition rate and variability in usage, it seems likely that students used the system in many different ways, perhaps ways different than those anticipated by the researcher.

The sensitivity of the levels is further confounded by the potential *covert* activities of the learner. The framework focuses on *overt* activities, but students might be engaging in a variety of *covert* activities (Chi, 2009). For example, during a passive reading activity, a student might spontaneously engage in self-explaining the material in his or her mind. In one study, while learning about the circulatory system, participants spent an average of an hour in a passive condition, with a range of twenty-two minutes to two hours, forty-seven minutes (Chi, De Leeuw, Chiu, & Lavancher, 1994). Although the participants read the materials with an experimenter present, it is not clear why some participants spent over seven times longer reading the material than others. Perhaps the slow reading can in part be explained by the time taken to complete *covert* activities on the part of the participant.

Only one study, in which students learned about psychological concepts, attempts to use convergent measures to assess compliance with the given requirements of each condition. The study compares the advantages of generating questions and answers, summarizing the material, and note taking with review (King, 1992). Participants were observed both in terms of their overt actions and self-reported implementation of the study strategy. The researcher noted through observation that all participants did what was asked of them in terms of questioning and answering, summarizing, or taking notes and reviewing. Further, self-reports revealed that each participant properly implemented the self-questioning and summarizing techniques. Note taking with review was not included in the self-report assessment. The author does note, however, that students might have used previously learned strategies—strategies that could potentially conflict with the assumptions of the framework.

The current study attempts to address the issue of participant compliance with the necessary requirements of passive, active, and constructive activities. Compliance was assessed in two ways: through inferences based on what students created while learning, or observed activity, and through a questionnaire of what they did while learning, or reported activity. All participants rated their utilization of strategies common to all three levels of activities, not just the ones they were trained to use. These measures were designed to ensure—with a reasonable degree of confidence—that participants *were* doing the things they are supposed to and were *not* doing things they should not. In the case that participants were operating outside of their assigned activity level, the data could be reanalyzed such that it accounted for this deviance.

1.3.3 Detailed Learner Outcomes

In order to solve problems, students must acquire knowledge—and be able to demonstrate this knowledge—in a meaningful way (Mayer, 2002). Meaningful learning, as defined by Mayer, is that which goes beyond factual regurgitation. Successful learning can be demonstrated by understanding, applying, analyzing, evaluating, and creating domain-relevant information. Further, the ability to transfer knowledge to novel problems is critical to education (Barnett & Ceci, 2002). These types of mastery cannot be conveyed through an assessment resulting in a single, overall score. The assessment must contain items ranging from basic recall of factual information to novel transfer-type problems. Though it might be initially helpful to summarize a student's performance using a single score, further analysis must take a closer look at the different types of assessment items. A closer look should help reveal the specific learning benefits, as well as potentially shed light on the underlying cognitive processes.

In one study aimed at teaching students the concepts of maps, students placed stickers on a map to represent real-world objects. Researchers used error from the target (i.e., measuring the distance between the stickers and the ideal location on the maps) to determine learner performance (Kastens & Liben, 2007). Although this measurement offers a continuous variable through which to analyze performance, it fails to connect student error with potential causes. Students who self-explained showed a lower degree of error by placing the stickers closer to the correct location. They were able to locate items in the real world and likely match them with the objects on the map, resulting in more accurate placement. But, this measure does not help to explain why self-explanations yield better map comprehension.

There are a few studies that assess learning with a variety of assessment items. After learning about algebraic models of real-world problems, participants selected self-explanations from a list, generated their own explanations, or were transitioned from selecting to generating using a scaffold method. Students then completed assessments involving generating models, solving problems, telling the meaning of symbols, and answering transfer problems (Corbett, Wagner, Lesgold, Ulrich, & Stevens, 2006). Although the researchers found that all conditions performed equally well on problems similar to those used during training, the participants who generated their own explanations did better on transfer problems. It is possible that this constructive study technique—self-explanations—engaged cognitive processes different than those experienced by other participants, thus allowing them to solve these problems. Other researchers have shown interactions by using near and far transfer problems after learners explained or summarized (Coleman, Brown, & Rivkin, 1997), as well as recall and synthesis questions after taking notes, taking notes and reviewing, or reviewing someone else's notes (Kiewra et al., 1991). These types of varied assessments are useful when speculating about the potential underlying cognitive processes in the different learning conditions.

Assessments have been developed using strict rules in order to clearly differentiate levels of items. In Chi et al. (1994), participants read material relating to the circulatory system one sentence at a time. Participants in the experimental condition self-explained after each sentence they read. Items at four different levels were developed to assess learner knowledge. The most basic items, verbatim, could be answered by relying on information from only one sentence that was previously read. Comprehension-inference items required the learner to integrate information across two lines of text. Knowledge-inference required the learner to integrate information from multiple lines of text while incorporating prior knowledge. Finally, a fourth level assessed student understanding of system-wide ideas. This assessment provides a method of understanding learning outcomes by aligning the study method and potential underlying cognitive processes (i.e., self-explaining and creating new connections within the material) with the assessment items themselves.

The tests in the present study comprise three levels of assessment items. The basic level—factual recall—required information from only one sentence within the text to answer. Intermediate items—system knowledge—required the knowledge from multiple sentences, typically an entire paragraph. With both of these levels, information necessary to respond to the item is explicitly provided in the text. Finally, advanced items—inference making—required knowledge from multiple paragraphs to answer, in addition to making an inference that goes beyond the information presented in the text. Example items at each of these three levels are provided in Table 3.

Table 3
Example assessment items based on the text from Figure 1.

Assessment level	Example items
Factual recall	The use of biotic and abiotic resources is known as a species' ______.
System knowledge	Describe why a change must occur if two species with identical niches are both to survive. Describe the change that would occur.
Inference ability	Using the concept of ecological niches, describe an experiment in which you attempt to determine if two species have identical niches. Describe the changes, or lack of changes, you would expect to find if the species do or do not have identical niches.

1.3.4 Delayed Assessment

The ability to retain knowledge is just as important as the ability to acquire knowledge in the first place (Schmidt & Bjork, 1992). Despite this, only one study analyzed in Chi's original comparison includes a delayed test. After learning about various social science concepts, students were tested immediately after learning and also tested one week later (King, 1992). The results indicate that summarizing and self-questioning are both helpful for short-term learning, while only self-questioning yields longer-term results. This approach adds further clarity to the potential underlying cognitive processes since it might be found that although certain activities yield similar immediate results, they might not result in similar learning outcomes over time. The present study also employs delayed testing with a delay of one week.

1.3.5 Time-on-task

Time-on-task is an important variable in explaining achievement differences. Increasing the amount of time on learning a particular task has been shown to predict achievement increases (Stallings, 1980). Merely increasing the amount of time spent in an ineffective learning environment, however, will not yield better learning (Anderson, 1981). Chi's (2009) framework necessitates nonequivalent learning times across learning activity levels. As each level subsumes the ones below it, it is assumed that the activities indicative of lower levels are present in higher levels. For example, in order for a student to engage in an effective self-explaining activities (i.e., a constructive activity), they must first read the material to be self-explained (i.e., a hallmark of a passive activity), and identify main ideas in the material that they will then self-explain (i.e., a hallmark of an active activity). Performing the activities of higher levels presumes that the learner spends more time with the material, assuming that the same amount of material is presented to all of the learners.

As presented here, time-on-task is an important consideration in this framework, yet many of the studies fail to take learning time into account. For example, some learners were given no time constraint during the learning activity, which was presumably ended at the discretion of the learner him- or herself (e.g., Coleman et al., 1997; Kastens & Liben, 2007). Other studies make no direct mention of learning time (e.g., Klahr & Nigam, 2004; Trafton & Trickett, 2001). It seems that interpretation of the results, though they generally support the framework, would be difficult given the importance of time-on-task and the apparent lack of control of it in these studies.

There are studies, however, that explicitly consider the variable of time, although interpretations of this variable seem to raise even more questions. In one study, some participants learned about the circulatory system either by reading the material twice while others read it once and self-explained. The first group of participants was required to read the material twice in an effort to equalize learning times. Even so, the passive condition yielded an average time of one hour, six minutes and the constructive condition two hours, five minutes (Chi et al., 1994). Participants in the constructive condition spent twice as much time with the material, potentially having twice the time to comprehend and integrate the material. The variable of time was not considered in the analysis as a potential predictor of the performance difference between the two conditions.

Time was also used in an experiment in which participants learned to tie knots by watching a video. In the passive condition, students watched the video from beginning to end, but they were able to do so as many times as they wished. In the active condition, participants were able to control the video by stopping, rewinding, pausing, and slowing. Participants ended the learning experience once they could tie each knot accurately. Learning time was used as the dependent variable showing that active learning required *less* time than passive learning (Schwan & Riempp, 2004). Since all participants were able to demonstrate mastery of the knots—as opposed to determining the time when mastery was reached—time warrants further elucidation. This particular study also violates the assumption that higher activity levels require more time, assuming that each activity level subsumes the ones below it. This violation might lead one to speculate that participants in the passive condition wasted quite a bit of time waiting for each video to run from start to finish, when perhaps they were merely interested in a certain portion of the video.

Finally, some of the experiments controlled time-on-task by equating it across conditions, which would violate the assumptions of Chi's framework. Students learned about psychology concepts by either contrasting the concepts and then attending a lecture, contrasting the concepts and contrasting more concepts instead of attending the lecture, or summarizing the concepts and then attending the lecture (Schwartz & Bransford, 1998). All of the participants spent exactly the same amount of time on learning, and it was found that contrasting and attending the lecture resulted in optimal learning. It might be the case, however, that the optimal condition provided better exposure to the material by avoiding redundant tasks, as might have been encountered in the other two conditions. In another experiment, participants viewed a variety of virtual 3D objects and then had to indicate which of them they had seen previously by selecting from various 2D views of the objects. All participants spent exactly the same amount of time learning. Participants who had actively rotated the 3D objects were able to recognize them faster, albeit only about 8% faster. All participants were able to accurately select the 2D views (James et al., 2002). Both of the aforementioned experiments would violate the assumption that higher levels of activities take more time to complete.

The present study aims to provide students with access to the same materials while using time-on-task as a predictor of learning outcomes. It is expected that constructive activities will take the most time and passive the least, with active somewhere in between. Time was measured and analyzed to determine its role in student outcome differences. As a pragmatic concern for teachers, time is valuable in the classroom. It is important that learning benefits of higher activity levels outweigh the increase in time to complete them.

1.4 Hypotheses

1.4.1 Summary of Experiment

The present study aims to test the framework by comparing three of its levels in a single experiment using the same materials across those three levels, while attempting to ensure compliance with the activity requirements, assessing learning using items of various levels, determining learning over time, and understanding the importance of time-on-task. The first aspect (i.e., simultaneous comparison) was implemented in the design of the experiment. Compliance with the requirements of the levels of passive, active, and constructive learning was necessary for confident interpretation—or potential reinterpretation—of results. Finally, detailed assessment items, delayed testing, and time-on-task each play important roles in the four primary hypotheses currently under investigation.

1.4.2 Hypothesis 1: Passive < Active < Constructive

A primary tenet of the framework is that active learning results in better outcomes than passive, and constructive results in better outcomes than active and passive. Immediately following the learning activity, overall assessments should demonstrate this pattern of results. After a delay of one-week, constructive should still do better than active, and active should still do better than passive in an assessment of overall learning.

If higher levels of learning do in fact engage more complex cognitive processes, student outcomes were expected to be better for higher levels of activities.

1.4.3 Hypothesis 2: Equivalent Outcomes for Equivalent Activity

Another primary tenet of the framework is activities of the same level should yield similar learning outcomes. Immediately following learning, learning outcomes for equivalent levels (e.g., highlighting, which is active, versus copy/paste note taking, which is also active) should be equivalent overall. After a week delay, the same pattern should still be found. If similar cognitive processes are being engaged despite different overt activities, student outcomes are expected to not be different. A failure to reject the null would confirm this hypothesis; however, as will be discussed later in the Results and Discussion section, the power in this experiment was likely insufficient to confidently confirm this hypothesis.

1.4.4 Hypothesis 3: Activity is Matched to Assessment Level

Another tenet of the framework is that higher levels of activity result in deeper processing, yielding superior outcomes on more challenging problems. Immediately following the learning activity, all conditions should perform equally well on the level one factual recall assessment items. Active and constructive should perform similar to each other, but better than passive, on level two system knowledge items. Finally, constructive should outperform both active and passive on level three inference making items. This pattern of results was expected to hold true for both immediate and delayed testing.

1.4.5 Hypothesis 4: Time-on-task Accounts for Performance

Although some studies have attempted to equate time-on-task, the framework necessitates that higher levels of activity require more time to complete, assuming that all students cover the same material. Based on pilot testing, time-on-task could increase by a factor of two and three over the passive conditions for active and constructive learning, respectively. It seems likely, then, that time-on-task should contribute to predicting the outcomes for both immediate and delayed testing. It was expected that when time-on-task is entered as a covariate in the analyses, it should explain the differences in performance between conditions.

1.5 Experiment Overview

This experiment tested three of the levels of Chi's (2009) framework. A summary of the conditions is provided here.

The passive-read condition required that the participants merely read the instructional text as if they were studying for a test.

The active-highlight condition required that the participants highlight the most important sentence in each paragraph as they progressed through the text. The literature suggests that directing learners to focus on the most important sentence in each paragraph is the minimal requirement for any meaningful learning to occur (Igo et al., 2005). Further, this direction seems in line with the framework since the participants are performing actions, but are not likely to be engaging in information creation activities.

The active-notes condition required that the participants find the most important sentence in each paragraph and copy-then-paste the text at the top of the screen. In essence, the participants developed notes that did not go beyond what was explicitly presented in the text.

The constructive-explain condition required that the participants read each page (typically comprising 2-4 paragraphs). explain its meaning in their own words, and describe how it connects to previous material. The directions are similar to those used in the literature (Chi et al., 1994). The length of material for each explanation is a bit longer than usual; however, some studies have required students to explain only after reading an entire story (Coleman et al., 1997). The length of text to be explained in this experiment fell within the range of what has been done previously and was thus deemed appropriate.

The constructive-question condition required that participants read each page and generate a "think type" question (King, 1992). These questions were expected to elicit answers that went beyond what was explicitly presented. Participants were asked to type the question and provide an ideal answer. To help participants get started, a list of suggested question starters was provided (see Table 4).

Table 4
Suggested question prompts provided in the constructive-question condition (King, 1992).

1. Explain why…
2. Explain how...
3. What is the main idea of…?
4. How would you use… to…?
5. What is a new example of…?
6. What do you think would happen if...?
7. What is the difference between… and...?
8. How are… and… similar?
9. What conclusions can you draw about...?
10. How does… affect...?
11. What are the strengths and weakness of…?
12. What is the best… and why?
13. How is… related to… that you read about earlier?

CHAPTER 2:

METHODS AND PROCEDURE

2.1 Participants

Participants were men and women eighteen years of age or older and were recruited from the Georgia Institute of Technology undergraduate population via Experimetrix. Upon arriving at the laboratory, participants were informed of their rights in accordance with the guidelines of the Institutional Review Board of the Georgia Institute of Technology. After signing the informed consent form, they were randomly assigned to one of five conditions. The experiment comprised two sessions spaced one week apart. During the first session, participants read the material while engaging in one of the five activities and were immediately tested. During the second session, participants returned to the laboratory approximately one week later to complete a delayed test. A summary of the sample size is shown in Table 5.

Table 5
Number of participants in each condition.

Condition	Initial session (N=55)	Delayed session (N=52)
Passive-read	10	10
Active-highlight	12	11
Active-notes	12	10
Constructive-explain	11	11
Constructive-question	10	10

2.2 Manipulations

The experiment comprised five conditions, one at the passive level and two each at the active and constructive levels. Detailed directions were provided by the computer to the participants describing what they were expected to do. These directions were given before the tutorial and before the instructional text. Abridged directions were provided on the top of every page as a reminder. The only additional item included in the detailed directions was a note encouraging the participants to read at their own pace as if studying for a test. The abridged directions provided to participants are shown in Table 6.

Table 6
Abridged directions for each of the five conditions.

Condition	Abridged directions to participant
Passive-read	Please read this page.
Active-highlight	Please read this page and... Highlight the most important sentence in each paragraph.
Active-notes	Please read this page and... Copy the most important sentence in each paragraph and paste it at the top of each page.
Constructive-explain	Please read this page and... Explain what each page means. Connect the most important information in each paragraph to what you already knew. Do not summarize each page. Describe in your own words why the text matters.
Constructive-question	Please read this page and... Develop a thought-provoking question for each page and an answer. Answers should include the most important information from each paragraph and something you already knew. Do not create "summary" questions. Questions and answers should address why the text matters.

2.3 Materials

The materials comprised the following: background questionnaire, pretest, tutorial, instructional text, immediate test, reported activity questionnaire, and delayed test.

2.3.1 Background Questionnaire

Although the participants were randomly assigned to the various conditions, potentially relevant background information was collected. The background questionnaire elicited from the participants their current number of accumulated credits, GPA, SAT scores, a list of the biology courses they had taken either in college or high school, as well as their major. This information was used to determine equivalence among the participants in each of the five assigned conditions.

2.3.2 Pretest

In a further effort to ensure initial equivalence among the participants, all participants answered a series of basic assessment items regarding community ecology. One item was asked for each of the eight subsections in the instructional text. All items were modified from their original multiple-choice format to free response. This test was used to determine the participants' pre-existing knowledge on the topic.

2.3.3 Tutorial

In order to familiarize participants with the learning environment, they read the directions and practiced the required study activity using two pages about chemistry. Upon completion of the tutorial, detailed feedback was provided to the participants to ensure that they were adequately adhering to the requirements of the condition. A rubric of the aspects of compliance is shown in Table 7.

Table 7
Rubric of requirements used to provide feedback to participants.

Condition	Requirements
Passive-read	(Since participants in this condition do not create any outputs, only the current time was recorded. This was done so that participants in all conditions would interact with the experimenter at the same points during the experiment.)
Active-highlight Active-notes	One highlight/note per paragraph Complete sentence highlighted/noted
Constructive-explain Constructive-question	At least one idea per paragraph Incorporated with explanation/question-answer Info from prior knowledge or previous page

2.3.4 Instructional Text

The instructional text was based on sections 3-5 of chapter 54 from the textbook *Biology* (Campbell & Reece, 2007). The textbook was selected based on strong recommendation by biology faculty at the Georgia Institute of Technology. The chapter on community ecology was selected because of its focus on systems, rather than on a list of declarative facts. Minor changes were made to the text importing it from its original textbook-formatted version to a computer-based interface. Figures were placed directly after the paragraph in which they are referenced. References to specific chapter numbers were removed. Also removed were review questions and inquiry-based sidebars. Finally, section summaries were removed and any other within-text questions were changed to statements to avoid encouraging participants to engage in actions similar to those being manipulated. Otherwise, the bulk of the text remained unchanged.

The participants interacted with the text through an interface shown in Figure 3. Area one was where abridged directions were displayed. These directions were an abridged version of the ones seen in the tutorial, serving as a reminder throughout the experiment. Area two was used in only some conditions (i.e., active-notes and both constructive conditions) where text was to be pasted or typed by the participant. It was left blank and not used in the passive-read and active-highlight conditions. Area three was where the instructional text and figures were shown.

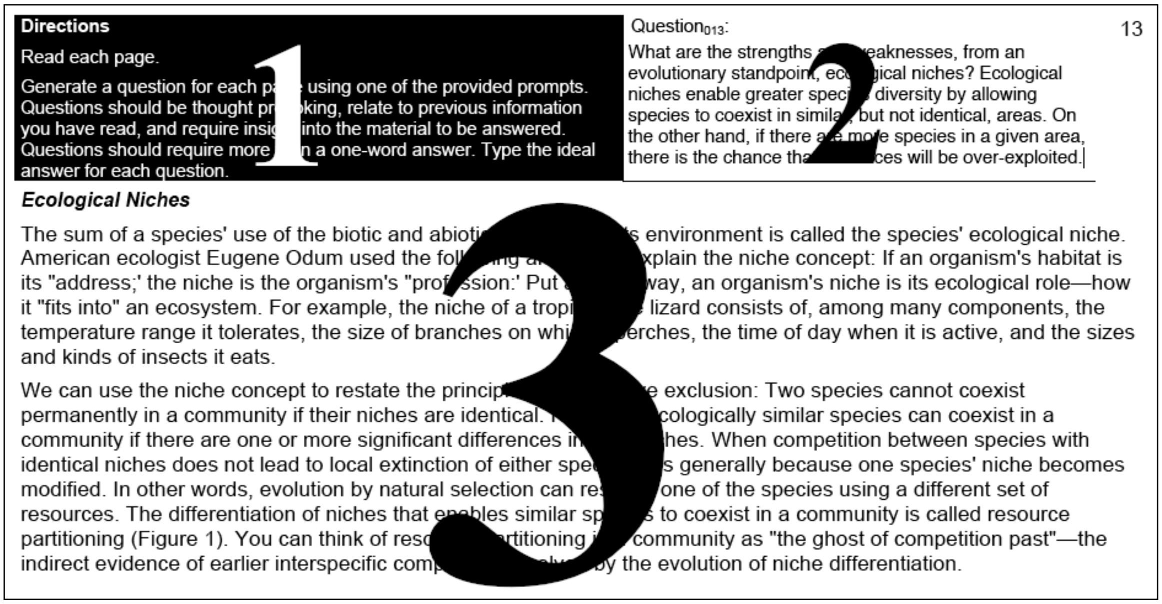

Figure 3. Example of the learning interface taken from the constructive-question condition.

2.3.5 Immediate and Delayed Tests

All assessment items were developed specifically for this experiment. Items were taken from all sections of the text. Multiple people, including a current second-year undergraduate biology student, reviewed the items for ambiguity and other potential issues. The items were piloted and adjusted for difficulty (i.e., easy questions were made more difficult, difficult questions were made easier). Example items and associated instructional text are provided in Figure 4.

2.3.6 Level One Items: Factual Recall

Assessment items at this level required the learner to remember verbatim information from the text. Items were developed representing the entire range of the text. Items were answerable with knowledge of a single sentence from the text and were written in a fill-in-the-blank format.

2.3.7 Level Two Items: System Knowledge

Assessment items at this second level, called system knowledge, also relied on verbatim information from the text. With these items, however, answers required knowledge of various aspects of a particular system described in the text. Items were designed such that multiple pieces of information from a single paragraph were necessary to fully answer.

2.3.8 Level Three Items: Inference Making

Inference making assessment items required that participants not only remember factual information and knowledge of systems, but also required participants to provide responses that went beyond what was explicitly presented in the text. Typically, the items were designed in the style of "what if." For example, participants were asked to determine the outcomes of a hypothetical experiment under the conditions that certain variables were altered.

Ecological Niches The sum of a species' use of the biotic and abiotic resources in its environment is called the species' ecological niche. American ecologist Eugene Odum used the following analogy to explain the niche concept: If an organism's habitat is its "address;' the niche is the organism's "profession:' Put another way, an organism's niche is its ecological role—how it "fits into" an ecosystem. For example, the niche of a tropical tree lizard consists of, among many components, the temperature range it tolerates, the size of branches on which it perches, the time of day when it is active, and the sizes and kinds of insects it eats. We can use the niche concept to restate the principle of competitive exclusion: Two species cannot coexist permanently in a community if their niches are identical. However, ecologically similar species can coexist in a community if there are one or more significant differences in their niches. When competition between species with identical niches does not lead to local extinction of either species, it is generally because one species' niche becomes modified. In other words, evolution by natural selection can result in one of the species using a different set of resources. The differentiation of niches that enables similar species to coexist in a community is called resource partitioning (Figure 1). You can think of resource partitioning in a community as "the ghost of competition past"—the indirect evidence of earlier interspecific competition resolved by the evolution of niche differentiation.	
Factual recall	The use of biotic and abiotic resources is known as a species' ______.
System knowledge	Describe why a change must occur if two species with identical niches are both to survive. Describe the change that would occur.
Inference making	Using the concept of ecological niches, describe an experiment in which you attempt to determine if two species have identical niches. Describe the changes, or lack of changes, you would expect to find if the species do or do not have identical niches.

Figure 4. Example assessment items with associated instructional text.

2.3.9 Reported Activity Questionnaire

In an effort to uncover any covert activities potentially utilized by participants that might go beyond what is requested of them, a questionnaire was developed based on the list of suggested actions for passive, active, and constructive activities (Chi, 2009). There were five statements regarding passive activities, five statements for active, and six statements for constructive. Participants were asked to rate the frequency with which they engaged in each of these activities either overtly or covertly on a 5-point Likert scale. The statements are shown in Table 8.

Table 8
Statements rated by participants regarding their activities during the learning task.

Passive statements
1. I read at about the same speed as I typically read.
2. I read faster than I typically read.*
3. I read slower than I typically read.
4. I re-read some of the text.
5. I skimmed sections of the text.*

Active statements
6. I highlighted portions of the text.**
7. I made notes of the text by copying-and-pasting.**
8. I identified the most important ideas.
9. I summarized or paraphrased the text.
10. I focused my attention on the text.

Constructive statements
11. I explained in my own words what the text meant to me.**
12. I developed thought-provoking questions and answers about the text.**
13. I connected the text to ideas I already knew.
14. I connected the text to ideas I had read on previous pages.
15. I made hypotheses or predictions about the text.
16. I justified or provided reasons why concepts in the text occur.

*These passive statements are "poor" passive activities; others are considered to be "good" passive activities.
**Statements are specific to a particular condition; all others could be performed regardless of assigned condition.

2.4 Procedure

Participants were assigned randomly to one of the five conditions. The experimenter described that the experiment takes place entirely on the computer. The experimenter encouraged the participants to work at a natural pace, as if they were studying for a test. The participants then completed the background questionnaire and pretest, after which the participants were instructed to stop and wait for the experimenter. Once all participants finished the pre-test, they each completed a short tutorial for their assigned study technique. The experimenter described the necessary requirements for the conditions and instructed the participants to start the tutorial. The experimenter reviewed their work for compliance with the requirements and feedback was given in accordance with Table 7. If the experimenter found that a participant did not adequately adhere to the guidelines of the condition, the participant was asked to fix whatever was wrong in the tutorial before moving on. In some cases, participants were not able to properly fix the issues but were still allowed to continue with the experiment.

Once all participants completed the tutorial and received approval from the experimenter, they began reading the instructional text, performing any necessary actions. The instructional text was divided into three main sections. After each of the sections, the participants were instructed to stop and call the experimenter. The experimenter briefly reviewed the completed work, provided any necessary feedback, and instructed the participants to continue. Once all of the text was read, the participants answered the immediate test assessment items and responded to the reported activity questionnaire. They returned approximately one week later to complete the delayed test.

CHAPTER 3:

RESULTS AND DISCUSSION

3.1 Participant Characteristics

The qualitative demographic data were coded in a quantitative manner so that they could be used in statistical analyses. Major was coded as "1" if the participant reported biology, biomedical engineering, or biochemistry, while all other majors were coded as "0." Courses were coded for the number of courses taken by the participant.

For time-on-task, the omnibus ANOVA of main effect of study condition was statistically significant, $F(4, 52) = 43.031$, $p < 0.001$. Tukey *post hoc* comparison of the five groups indicates that the reading ($M = 27.55$, $SD = 8.49$), highlighting ($M = 36.42$, $SD = 5.85$), and note taking ($M = 40.92$, $SD = 7.89$) participants took less time than the self-explanation ($M = 79.82$, $SD = 20.44$) and self-question participants ($M = 97.18$, $SD = 25.43$). For SAT quantitative scores, the omnibus ANOVA of main effect of study condition was statistically significant, $F(4, 50) = 3.825$, $p < 0.001$. Tukey *post hoc* comparison of the five groups indicates that the highlighting ($M = 648.33$, $SD = 46.87$) and self-explaining ($M = 636.36$, $SD = 92.77$) participants had lower scores than the note-taking ($M = 733.33$, $SD = 58.05$) participants.

Time-on-task and SAT quantitative scores were entered as covariates when analyzing test performances (see results below). For the remaining demographic variables (i.e., credits; GPA; SAT verbal; number of biology courses taken in high school; number of biology courses taken in college; and having declared a major of biology, biochemistry, or biomedical engineering) were not statistically significant to the criterion of $p < 0.05$ and were not used as covariates.

Descriptive statistics for participant performance are reported in Table 9 and Table 10 for the immediate test and delayed test, respectively. The number of participants, the overall score, and the score breakdown by level 1 (factual recall), 2 (system knowledge), and 3 (inference making) questions are shown.

Table 9
Estimated means (out of 1.00) and standard errors for immediate test, both overall and by level of questions with time-on-task and SAT quantitative as covariates.

			Questions by Level (1, 2, 3)		
Condition	n	Overall Score	Factual	System	Inference
Read	10	0.54 (0.078)	0.57 (0.085)	0.60 (0.097)	0.44 (0.090)
Highlight	12	0.46 (0.066)	0.50 (0.072)	0.54 (0.082)	0.35 (0.076)
Notes	12	0.44 (0.066)	0.47 (0.073)	0.49 (0.082)	0.36 (0.076)
Explain	11	0.36 (0.074)	0.42 (0.081)	0.31 (0.091)	0.32 (0.085)
Question	10	0.39 (0.096)	0.42 (0.105)	0.30 (0.119)	0.45 (0.110)

Table 10
Estimated means (out of 1.00) and standard errors for delayed test, both overall and by level of questions with time-on-task and SAT quantitative as covariates.

			Questions by Level (1, 2, 3)		
Condition	n	Overall Score	Verbatim	System	Inference
Read	10	0.48 (0.067)	0.65 (0.079)	0.46 (0.080)	0.34 (0.070)
Highlight	11	0.37 (0.058)	0.44 (0.068)	0.34 (0.070)	0.35 (0.061)
Notes	10	0.37 (0.060)	0.50 (0.070)	0.37 (0.072)	0.25 (0.063)
Explain	11	0.28 (0.061)	0.38 (0.072)	0.21 (0.073)	0.25 (0.064)
Question	10	0.32 (0.080)	0.40 (0.094)	0.26 (0.096)	0.29 (0.084)

3.2 Coding of Test Items

Responses to test items were coded independently by two raters using a previously piloted rubric. For the pretest, the coders had 99.1% initial agreement and resolved the remaining disagreements through discussion. For the immediate test, the coders had 86.1% initial agreement and resolved 13.8% of the disagreements through discussion. A third party broke the tie on the remaining disagreements comprising 0.1% of the codes. For the delayed test, the coders had 84.0% initial agreement and resolved 15.9% of the disagreements through discussion. A third party broke the tie on the remaining disagreements comprising 0.1% of the codes. Averages were calculated for each of the tests, resulting in scores between 0.00 and 1.00.

3.3 Hypothesized Learner Outcomes

The four *a priori* hypotheses of this experiment are discussed. A summary of the statistical tests is provided in Table 11.

3.3.1 Hypothesis 1: Passive < Active < Constructive

It was expected that those participants in higher levels of activity would do better overall. Omnibus ANOVAs returned non-significant results for both the immediate and delayed tests.

3.3.2 Hypothesis 2: Equivalent Outcomes for Equivalent Activities

It was expected that those participants in the same level of activity (e.g., highlighting and note-taking were both active activities) should have done similarly overall. Since the omnibus ANOVA investigating differences among the three levels was non-significant, no conclusions can be made about this hypothesis.

3.3.3 Hypothesis 3: Activity is Matched to Assessment Level

It was expected that participants in higher levels of activity would be able to perform better on more challenging questions. Omnibus ANOVAs for each of the three levels of questions for both the immediate and delayed tests were non-significant.

3.3.4 Hypothesis 4: Time-on-task Accounts for Performance

It was expected that controlling for time-on-task would explain some of the differences among conditions. In only one case (i.e., level 2 questions on the immediate test) did the ANCOVA yield a significant result. Although using time-on-task as a covariate typically increased the *F*-values for most of the ANOVAs and ANCOVAs, no conclusion can be made about the significance of time-on-task.

Table 11
A summary of ANCOVA results for the four hypotheses.

Test	Questions	By Condition	By Condition with Time-on-Task as Covariate
Immediate	Overall	F(4,48) = .63, p > .05	F(1,48) = 1.89, p > .05
	Level 1: Factual	F(4,48) = .40, p > .05	F(1,48) = 1.93, p > .05
	Level 2: System	F(4,48) = .90, p > .05	F(1,48) = 4.98, **p < .05**
	Level 3: Inference	F(4,48) = .70, p > .05	F(1,48) = .02, p > .05
Delayed	Overall	F(4,45) = 1.11, p > .05	F(1,48) = 2.42, p > .05
	Level 1: Factual	F(4,45) = 1.91, p > .05	F(1,48) = 2.18, p > .05
	Level 2: System	F(4,45) = 1.09, p > .05	F(1,48) = 2.66, p > .05
	Level 3: Inference	F(4,45) = .62, p > .05	F(1,48) = 1.02, p > .05

3.3.5 Sample Size in Current Experiment

One explanation for the lack of significant results might lie in the limited sample size. Assuming a medium effect size of 0.25—which is inline with other similar experiments (Atkinson, Renkl, & Merrill, 2003)—to achieve a power of 0.95, a sample size of approximately sixty participants per condition would be necessary to yield statistical significance. This number of participants is substantially more than the actual number of participants in this experiment. Upon reviewing the relevant literature Chi (2009) used to support the framework, many studies used approximately the same number of participants per conditions as this experiment. Typically, conditions comprised nine-twenty participants each (e.g., Coleman et al., 1997; James et al., 2002; Kiewra et al., 1991; King, 1992; Mitrovic, 2005; Schwan & Riempp, 2004; Schwartz & Bransford, 1998). Of particular note is that Chi's (1994) study comparing reading and self-explanations—a very similar experiment to the current one—utilized about twelve participants per conditions.

3.4 Variables Potentially Accounting for Performance Differences

3.4.1 Observed Activity Does Not Predict Performance

A variable, observed activity, was coded based on the output from the participants during the learning experiences (e.g., their notes, their self-explanations). Although participants were trained on and received feedback regarding their assigned study technique, participants were not always able to properly complete the activity according to the rubric in Table 7. Participants were assigned a score of one, two or three for each page of the nineteen pages of the learning activity. Everyone received one point for reading the page. Participants in the active and constructive conditions could earn an additional point for identifying main ideas from each paragraph. Finally, participants in the constructive conditions could earn an additional point for including information not explicitly provided in the text. Two independent raters coded all pages. Initial agreement was 84.7% and 15.2% were resolved through discussion. A third party broke the tie on the remaining disagreements comprising 0.1% of the codes. The scores for each of the nineteen pages were averaged for each participant. Each participant received a score between one and three for their observed activity variable. This variable does not enter in any of the subsequent regressions as a reliable predictor of performance.

3.4.2 Self-reported Activities Inconsistent with Assigned Activities

All participants responded to a series of statements in terms of the frequency with which they engaged in various activities, either overtly on the computer or covertly in their mind. Four of the statements were directly applicable to specific conditions. For example, only participants in the note-taking condition should have responded positively to the statement "I made notes of the text by copying-and-pasting." There were similar statements for the highlight, self-explanation, and self-questioning conditions (see Table 8 statements with **). Using a chi-square test, it was only among these four statements that participants from the five conditions responded differently (highlighting-specific statement: X^2 (16, N = 57) = 44.87, $p < 0.01$; note-taking-specific statement: X^2(16, N = 57) = 62.18, $p < 0.01$; self-explaining-specific statement: X^2(16, N = 57) = 26.69, $p = 0.045$; self-questioning-specific statement: X^2(16, N = 57) = 48.28, $p < 0.01$). These results make it clear that people did do what was asked of them, but they also did other behaviors. For all other statements, participants did not respond differently to the passive, active, and constructive statements *regardless of the condition to which they were assigned*. Chi (2009) states that a caveat to the framework involves overt and covert activities. Although the framework is operationalized in terms of overt activities, students might engage in covert activities not specific to the condition in which they are supposed to be studying. A passive reader could self-explain in her head. Verbatim note takers could randomly select text without thinking about what they are doing in a passive manner. Participants in this experiment spontaneously engaged in a variety of study techniques that went beyond what they were asked to do.

The self-report responses regarding what participants believed they did while learning were summed. From the statements in Table 8, two were removed from each of the passive, active, and constructive statements (i.e., those with either a single or double asterisk). Since the intention of the summed score was to indicate to what degree participants *did* engage, statements two and five were removed as they were deemed "negative" passive behaviors. Statements six, seven, eleven, and twelve were removed because they represented behaviors specific to particular conditions. Table 12 shows how the remaining responses were used to form the reported activity scores based on participants' self-reported behaviors.

Table 12
Procedure for determining scores based on self-reports by participants.

Variable	Range of values	Calculated by summing responses (1-5) to these statements:
Reported-passive	3-15	I read at about the same speed as I typically read. I read slower than I typically read. I re-read some of the text.
Reported-active	3-15	I identified the most important ideas. I summarized or paraphrased the text. I focused my attention on the text.
Reported-constructive	4-20	I connected the text to ideas I already knew. I connected the text to ideas I had read on previous pages. I made hypotheses or predictions about the text. I justified or provided reasons why concepts in the text occur.

3.4.3 GPA, SAT Verbal, and Self-reported Activities Predict Performance

A step-wise regression approach was used to investigate the variables that predict performance. The potential predictors were time-on-task; credits earned; GPA; SAT verbal; SAT quantitative; number of biology courses taken in high school; number of biology courses taken in college; having declared a major of biology, biochemistry, or biomedical engineering; observed activity; reported-passive/active/constructive; and pre-test scores. Dummy-coding was used for the categorical predictor of assigned condition. The variables were entered stepwise with the criteria set for *F*-to-enter at 0.05 and *F*-to-remove at 0.10. Four regressions were run for the immediate test: overall, level one questions, level two questions, and level three questions. For a summary of the results, see Table 13, Table 14, Table 15, and Table 16.

Table 13
Regression model for predicting overall immediate test performance.

Variable	*B*	*SE B*	β	R^2 Total	Correlation
Reported-active	.023	.009	.269	.198	.445**
SAT verbal	.001	.000	.385	.330	.440**
GPA	.108	.034	.327	.437	.403**
Reported-constructive	.012	.005	.241	.490	.283*

Correlations significant to $*p < .05$ $**p < .01$

Table 14
Regression model for predicting level one question immediate test performance.

Variable	*B*	*SE B*	β	R^2 Total	Correlation
Reported-active	.033	.011	.346	.219	.468**
SAT verbal	.001	.000	.292	.280	.330**
Reported-constructive	.014	.006	.252	.337	.313*

Correlations significant to $*p < .05$ $**p < .01$

Table 15
Regression model for predicting level two question immediate test performance.

Variable	*B*	*SE B*	β	R^2 Total	Correlation
GPA	.139	.047	.334	.170	.412**
Biology Major	.140	.072	.228	.267	.376**
SAT verbal	.001	.000	.327	.326	.326**
Time-on-task	.002	.001	.241	.377	.195

Correlations significant to $*p < .05$ $**p < .01$

Table 16
Regression model for predicting level three question immediate test performance.

Variable	B	$SE\ B$	β	R^2 Total	Correlation
SAT verbal	.001	.000	.416	.236	.485**
GPA	.097	.042	.257	.314	.325**
Reported-active	.024	.011	.244	.371	.352**

Correlations significant to *$p < .05$ **$p < .01$

Delayed tests were investigated using four additional step-wise regressions. The respective immediate test scores were included as an additional predictor. For example, for the regression on performance with level two questions for the delayed test, the performance on level two questions for the immediate test was used as a predictor. A summary of the results is found in Table 17, Table 18, Table 19, and Table 20.

Table 17
Regression model for predicting overall delayed test performance.

Variable	B	$SE\ B$	β	R^2 Total	Correlation
Immediate test overall performance	.770	.068	.848	.686	.828**
Credits	.001	.000	.197	.725	.113

Correlations significant to *$p < .05$ **$p < .01$

Table 18
Regression model for predicting level one question delayed test performance.

Variable	B	$SE\ B$	β	R^2 Total	Correlation
Immediate test level 1 performance	.582	.098	.610	.453	.673**
GPA	.086	.035	.252	.513	.406**

Correlations significant to *$p < .05$ **$p < .01$

Table 19
Regression model for predicting level two question delayed test performance.

Variable	B	$SE\ B$	β	R^2 Total	Correlation
Immediate test level 2 performance	.575	.090	.670	.449	.670**

Correlations significant to *$p < .05$ **$p < .01$

Table 20
Regression model for predicting level three question delayed test performance.

Variable	B	$SE\ B$	β	R^2 Total	Correlation
Immediate test level 3 performance	.352	.100	.430	.262	.512**
GPA	.083	.036	.278	.333	.405**

Correlations significant to *$p < .05$ **$p < .01$

In an effort to determine which variables actually predicted performance, GPA and SAT verbal scores were quite consistently significant. This was not unexpected. GPA is perhaps a measure of competitiveness on school-related tasks, and the task in this experiment was exactly that. Further, given the intense focus on reading, verbal ability is likely to help one do well. What is more interesting is that participants' self-reports of their behaviors more readily predicted performance than what they were told to do. Reporting active and constructive activities was highly correlated with better overall and level one performance. In terms of delay, performance on the immediate test was the best predictor. Generally, those participants who reported higher levels of activity did better than those who did not, despite what they were told to do.

3.4.5 Course of Study Predicts Self-reported Activities

A step-wise regression approach was used to investigate the variables that might predict reported activities (i.e., reported-passive/active/constructive). The potential predictors were time-on-task; credits earned; GPA; SAT verbal; SAT quantitative;

number of biology courses taken in high school; number of biology courses taken in college; having declared a major of biology, biochemistry, or biomedical engineering; observed activity; and pre-test scores. Dummy-coding was used for the categorical predictor of assigned condition. The variables were entered stepwise with the criteria set for *F*-to-enter at 0.05 and *F*-to-remove at 0.10. Three regressions were run: reported-passive, reported-active, reported-constructive (see Table 21, Table 22, and Table 23).

Table 21
Regression model for predicting reported-passive.

Variable	*B*	*SE B*	β	R^2 Total	Correlation
Biology major	1.613	.577	.356	.100	.317**
Self-questioning condition	-1.244	.577	-.28275	.174	-.223

Correlations significant to *$p < .05$ **$p < .01$

Table 22
Regression model for predicting reported-active.

Variable	*B*	*SE B*	β	R^2 Total	Correlation
Biology major	1.852	.706	.327	.134	.367**
Self-questioning condition	1.566	.706	.276	.209	.323**

Correlations significant to *$p < .05$ **$p < .01$

Table 23
Regression model for predicting reported-constructive.

Variable	*B*	*SE B*	β	R^2 Total	Correlation
Highlighting condition	-4.266	1.186	-.457	.148	-.385**
Note-taking condition	-2.392	1.186	-.259	.210	-.170

Correlations significant to *$p < .05$ **$p < .01$

As a final exploration, self-report scores were investigated. Assigned activity surfaced as significant predictors in each of the three reported activity variables. For reported-passive and reported-active, being assigned to the questioning condition resulted in a negative and positive correlation, respectively. It seems that questioning prevented good passive learning, to some degree. It encouraged, however, active learning. These participants seemed to have spent more time trying to find main ideas than carefully reading the text. Being assigned to highlighting or note-taking negatively correlated to being reported-constructive. These participants were likely focused on completing the task of finding the main ideas rather than trying to make connections. Further investigation of these predictors is warranted to more fully understand these complexities.

It was found that, generally, those participants who reported higher activity were more likely to have declared biology as a major. This is interesting for a number of reasons. Given that students select a course of study that they enjoy, perhaps the material was more interesting to biology majors, thus leading them to have higher levels of activity. Another explanation might be that since biology majors have high levels of exposure to reading similar types of material, they have learned—and were able to apply more readily—study techniques suited to learning biology text.

CHAPTER 4:

CONCLUSION

The current study was designed to confirm Chi's (2009) framework. Experimental manipulations were carefully controlled to ensure that all participants had the same opportunity to learn, and that the primary difference between the conditions was the directions given to the participants in how to engage with the material. The general hypotheses of the framework were not supported, even when time-on-task was controlled for. Although the sample size was small, it was similar to other studies used to support the framework. The framework was created to help designers of educational experiences tap into certain cognitive processes through various overt activities. Based on the participants in this study, the framework is not directly supported.

There were variables, however, that appeared to account for performance differences. Despite what they are asked to do, learners might bring to bear outside experiences while studying. In this experiment, participants were able to self-report behaviors that predicted their performance. These behaviors were aligned to Chi's framework of passive, active, and constructive. Based on self-report of activities—and not what they were asked to do by the experimenter—engaging in higher levels of active and constructive learning yields better learning. Assigning students to study in a particular manner does not seem to predict performance, but presumably activity as described by the passive-active-constructive framework can yield predictable learning.

More work is necessary in understanding the variables that contribute to a learner's willingness to engage in the material with assigned or previously learned study techniques. Students in this experiment were selected from a population representing some of the brightest undergraduate students in the country. Given this restriction of range issue, repeating this experiment in a setting with students at the other end of the educational spectrum would be informative. Perhaps these students might have less rigidly developed study habits, thus allowing the assigned treatment to influence their behavior more readily. In this case, perhaps, the *a priori* hypotheses will be supported.

REFERENCES

Anderson, L. W. (1981). Instruction and Time-on- Task: a Review. *Journal of Curriculum Studies, 13*(4), 289 - 303.

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning From Studying Examples to Solving Problems: Effects of Self-Explanation Prompts and Fading Worked-Out Steps. *Journal of Educational Psychology, 95*(4), 774-783. doi: 10.1037/0022-0663.95.4.774

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128*, 612-637.

Campbell, N. A., & Reece, J. B. (2007). *Biology*. Boston, MA: Benjamin Cummings.

Chi, M. T. H. (2009). Active-Constructive-Interactive: A Conceptual Framework for Differentiating Learning Activities. *Topics in Cognitive Science, 1*, 73-105.

Chi, M. T. H., De Leeuw, N., Chiu, M.-H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*(3), 439-477.

Coleman, E. B., Brown, A. L., & Rivkin, I. D. (1997). The Effect of Instructional Explanations on Learning From Scientific Texts. *Journal of the Learning Sciences, 6*(4), 347 - 365.

Corbett, A., Wagner, A., Lesgold, S., Ulrich, H., & Stevens, S. (2006). *The impact on learning of generating vs. selecting descriptions in analyzing algebra example solutions*. Paper presented at the Proceedings of the 7th international conference on Learning sciences, Bloomington, Indiana.

Igo, L. B., Bruning, R., & McCrudden, M. T. (2005). Exploring Differences in Students' Copy-and-Paste Decision Making and Processing: A Mixed-Methods Study. *Journal of Educational Psychology, 97*(1), 103-116.

James, K. H., Humphrey, G. K., Vilis, T., Corrie, B., Baddour, R., & Goodale, M. A. (2002). “Active” and “passive” learning of three-dimensional object structure within an immersive virtual reality environment. *Behavior Research Methods, Instruments, & Computers, 34*(3), 383-390.

Kastens, K. A., & Liben, L. S. (2007). Eliciting Self-Explanations Improves Children's Performance on a Field-Based Map Skills Task. *Cognition and Instruction, 25*(1), 45 - 74.

Kiewra, K. A., DuBois, N. F., Christian, D., McShane, A., Meyerhoffer, M., & Roskelley, D. (1991). Note-taking functions and techniques. *Journal of Educational Psychology, 83*(2), 240-245. doi: 10.1037/0022-0663.83.2.240

King, A. (1992). Comparison of Self-Questioning, Summarizing, and Notetaking-Review as Strategies for Learning From Lectures. *American Educational Research Journal, 29*(2), 303-323. doi: 10.3102/00028312029002303

Klahr, D., & Nigam, M. (2004). The Equivalence of Learning Paths in Early Science Instruction. *Psychological Science, 15*(10), 661-667.

Mayer, R. E. (2002). Rote versus Meaningful Learning. *Theory into Practice, 41*(4), 226-232.

Mitrovic, A. (2005). Scaffolding answer explanation in a data normalization tutor. *Facta universitatis - Series: Electronics and Energetics, 18*(2), 151-163.

Schmidt, R. A., & Bjork, R. A. (1992). New Conceptualizations of Practice: Common Principles in Three Paradigms Suggest New Concepts for Training. *Psychological Science, 3*(4), 207-217.

Schwan, S., & Riempp, R. (2004). The cognitive benefits of interactive videos: learning to tie nautical knots. *Learning and Instruction, 14*(3), 293-305.

Schwartz, D. L., & Bransford, J. D. (1998). A Time For Telling. *Cognition and Instruction, 16*(4), 475 - 5223.

Smallwood, J., McSpadden, M., & Schooler, J. W. (2008). When attention matters: The curious incident of the wandering mind. *Memory & Cognition, 36*, 1144-1150.

Stallings, J. (1980). Allocated Academic Learning Time Revisited, or Beyond Time on Task. *Educational Researcher, 9*(11), 11-16. doi: 10.3102/0013189x009011011

Trafton, J. G., & Trickett, S. B. (2001). Note-Taking for Self-Explanation and Problem Solving. *Human-Computer Interaction, 16*(1), 1 - 38.